

**MODELO DE REGRESIÓN LOGÍSTICA COMO ALTERNATIVA PARA MEDIR LA
PROBABILIDAD DE DESERCIÓN TEMPRANA EN LA UNIVERSIDAD DE LOS
LLANOS PERIODO 2015-2 – 2018-1**

AUTORES

**CRISTIAN ARLEY ARISMENDY FUENTES
NATHALIA LORENA MORALES PARRADO**

DIRECTOR

ERNESTO LEONEL CHÁVEZ HERNÁNDEZ

**UNIVERSIDAD DE LOS LLANOS
FACULTAD DE CIENCIAS ECONÓMICAS
ECONOMÍA
2018**

**MODELO DE REGRESIÓN LOGÍSTICA COMO ALTERNATIVA PARA MEDIR LA
PROBABILIDAD DE DESERCIÓN TEMPRANA EN LA UNIVERSIDAD DE LOS
LLANOS PERIODO 2015-2 – 2018-1**

COMITÉ DE PROGRAMA DE ECONOMÍA

**CRISTIAN ARLEY ARISMENDY FUENTES
NATHALIA LORENA MORALES PARRADO**

INFORME FINAL PARA OPTAR AL TÍTULO DE ECONOMISTA

**UNIVERSIDAD DE LOS LLANOS
FACULTAD DE CIENCIAS ECONÓMICAS
ECONOMÍA
2018**

AUTORIDADES ACADÉMICAS

PABLO EMILIO CRUZ CASALLAS
Rector (E).

DORIS CONSUELO PULIDO DE GONZÁLES
Vicerrectora Académica

GIOVANNY QUINTERO REYES
Secretario General

RAFAEL OSPINA INFANTE
Decano Facultad de Ciencias Económicas

ERNESTO LEONEL CHÁVEZ HERNÁNDEZ
Director programa de Economía

Nota de aceptación

Ernesto Leonel Chávez Hernández
Director programa de Economía

Ernesto Leonel Chávez Hernández
Director trabajo de grado

Villavicencio, septiembre 2018

DEDICATORIA,

A Ana, por ser madre, amiga y guía

A Lorena, por ser paciente y motivarme siempre

A Shirley por estar al tanto de todo

A todos aquellos que saben que aportaron en algún momento para que esto fuera posible.

Cristian Arismendy Fuentes

A Cristian por su apoyo, comprensión y enseñanza.

Nathalia Lorena Morales Parrado

TABLA DE CONTENIDO

1. RESUMEN.....	7
2. CUMPLIMIENTO DE OBJETIVOS.....	7
2.1. Cumplimiento del objetivo general	7
2.2. Cumplimiento de los objetivos específicos	8
3. DESCRIPCIÓN DE RESULTADOS/PRODUCTOS ENTREGABLES	10
4. CUMPLIMIENTO DE LA METODOLOGÍA.....	11
5. SEGUIMIENTO DEL CRONOGRAMA DE ACTIVIDADES A LA FECHA.	11
6. IMPACTO SOCIAL	12
7. DISCUSIÓN Y ANÁLISIS	13
Índice de deserción temprana por Facultad	13
Estimación del modelo logit	14
Relaciones de probabilidad.....	16
Bondad de ajuste del modelo.....	18
CONCLUSIONES	20
SIGLAS Y ABREVIATURAS	20
REFERENCIAS BIBLIOGRÁFICAS	21
LISTA DE ANEXOS	22

1. RESUMEN

Esta investigación tiene como principal propósito analizar la probabilidad de deserción temprana de los estudiantes de la Universidad de los Llanos mediante un modelo de regresión logística, además pretende identificar cuáles son las variables de mayor incidencia en la deserción temprana. Para la construcción del modelo se tomó como muestra la información de 574 estudiantes que ingresaron en la cohorte 2015-2 y que para el periodo 2018-1 eran registrados como vigentes o reportaban como último periodo matriculado cualquiera de los cuatro primeros semestres. Los resultados que arrojó esta investigación dan cuenta de que un buen puntaje en la prueba Saber 11, el ser mujer, no haber reprobado años durante el bachillerato, el haber cursado estudios antes y si los padres conviven disminuyen la probabilidad de deserción temprana; asimismo, el haber egresado de un colegio privado aumenta la probabilidad de deserción. De todas las facultades en las que se puede matricular el estudiante, las que mayor riesgo tienen de deserción temprana son la de Ciencias Básicas e Ingeniería y la Facultad de Ciencias de la Salud.

2. CUMPLIMIENTO DE OBJETIVOS

2.1. Cumplimiento del objetivo general

OBJETIVO GENERAL:	Analizar la probabilidad de deserción temprana de los estudiantes que ingresaron a la Universidad de los Llanos en la cohorte 2015-2 mediante un modelo de regresión logística para hacer recomendaciones de política educativa.		% de cumplimiento:	100%
RESULTADO OBTENIDO	ANEXO SOPORTE DEL DESARROLLO Y OBTENCIÓN DE RESULTADOS	DIFICULTADES	OBSERVACIONES	
Se logró construir un modelo de regresión logística a partir de la información de 574 estudiantes que ingresaron en la cohorte 2015-2. El modelo clasifica correctamente 54% de los retiros definitivos y el 74% de los	Anexo 1. manuscrito de artículo científico sometido para publicación	Se encontró que mucha de la información que reportan los estudiantes es falsa, asimismo, se identificó que muchos de los estudiantes que registraron retiro definitivo, en el	Se clasificaron los estudiantes como desertores a aquellos que presentaron su retiro definitivo o registraron como último semestre matriculado cualquiera de los cuatro primeros. Esto se hizo, debido a que aquellos que se retiran temporalmente tienen un	

vigentes. Para un acierto global del 67%.		periodo 2018-1 aún registraban como vigentes o en retiro temporal. También se encontró problemas con la poca cantidad de información que reportan los estudiantes, dado que hubo que excluir muchas observaciones por ausencia de datos.	plazo máximo de dos periodos consecutivos para renovar matrícula, en caso de no hacerlo se consideran desertores.
---	--	--	---

2.2. Cumplimiento de los objetivos específicos

OBJETIVO ESPECÍFICO:	Identificar las variables que mayor incidencia tienen en la deserción temprana de los estudiantes de la Universidad de los Llanos		% de cumplimiento:	100%
RESULTADO OBTENIDO	PRODUCTO (si aplica)	ANEXO SOPORTE DEL DESARROLLO Y OBTENCIÓN DE RESULTADOS	OBSERVACIONES	
Del análisis de 13 variables que se consideraron podrían explicar el fenómeno de la deserción temprana. Resultaron estadísticamente significativas 7 de ellas. Se encontró que un buen puntaje en la prueba Saber 11, el ser mujer, no haber reprobado años durante el bachillerato, el haber cursado estudios antes y si los padres conviven disminuyen la probabilidad de deserción temprana; asimismo, el haber egresado de un colegio privado aumenta la probabilidad de deserción. De todas las	Anexo 1. manuscrito de artículo científico sometido para publicación	Anexo 1. manuscrito de artículo científico sometido para publicación	Hubo que eliminar muchas observaciones debido a la ausencia de la información. Algunas variables no se tuvieron en cuenta debido a que la información presentaba inconsistencias.	

facultades en las que se puede matricular el estudiante, las que mayor riesgo tienen de deserción temprana son la de Ciencias Básicas e Ingeniería y la Facultad de Ciencias de la Salud.			
---	--	--	--

OBJETIVO ESPECÍFICO:	Analizar la probabilidad de permanencia de los estudiantes que se matriculan en la universidad de los Llanos.		% de cumplimiento:	100%
RESULTADO OBTENIDO	PRODUCTO (si aplica)	ANEXO SOPORTE DEL DESARROLLO Y OBTENCIÓN DE RESULTADOS	OBSERVACIONES	
Con el modelo se logró identificar las variables de mayor incidencia tienen para que un estudiante permanezca vigente en la Universidad de los Llanos. Además, el modelo clasificó correctamente el 74% de los estudiantes que se encuentran vigentes. Del mismo modo, se consiguieron las relaciones de probabilidad entre las variables y se construyeron perfiles de ingreso para futuros estudiantes.	Anexo 1. manuscrito de artículo científico sometido para publicación	Anexo 1. manuscrito de artículo científico sometido para publicación	Ninguna	

OBJETIVO ESPECÍFICO:	Recomendar políticas educativas enfocadas a prevenir la deserción temprana en la Universidad de los Llanos.		% de cumplimiento:	100%
RESULTADO OBTENIDO	PRODUCTO (si aplica)	ANEXO SOPORTE DEL DESARROLLO Y OBTENCIÓN DE RESULTADOS	OBSERVACIONES	
Con la identificación de las variables que mayor peso tienen para que un estudiante deserte de la Universidad de los Llanos, se logró formular algunas recomendaciones que se deben tener en cuenta para que desde las diferentes oficinas de la Universidad de los Llanos intervengan para prevenir que un estudiante abandone sus estudios	Manuscrito de artículo sometido para publicación.	Manuscrito de artículo sometido para publicación.	Ninguna	

3. DESCRIPCIÓN DE RESULTADOS/PRODUCTOS ENTREGABLES

RESULTADOS/PRODUCTOS ENTREGABLES	INDICADOR DE CUMPLIMIENTO	DESCRIPCIÓN DEL RESULTADO/ PRODUCTOS ENTREGABLES OBTENIDO	ANEXO SOPORTE
Elaboración, entrega y discusión de artículo publicable del trabajo de grado.	Compromiso alcanzado	Se elaboró un artículo científico el cual será sometido a aprobación por el Centro de Investigaciones de la FCE para dictaminar si es publicable o no	Manuscrito de artículo sometido para publicación
Socialización de los resultados del trabajo con la comunidad académica.	Compromiso pendiente	A partir de la culminación de la investigación, se procede a socializar los resultados con la comunidad académica	Fotografías y memorias del evento

4. CUMPLIMIENTO DE LA METODOLOGÍA

Hubo necesidad de verificar la base de datos, debido a que se encontró que algunos estudiantes que se habían retirado de la Universidad aún registraban como vigentes o en retiro temporal, lo cual provocó que las proporciones de estudiantes vigentes y desertores variaran de manera significativa. Además, fue necesario no incluir información concerniente a la dependencia económica y nivel de educación de la madre debido a que la información no era coherente. Asimismo, se eliminó la información de carácter territorial debido a que para esta cohorte de ingreso ningún estudiante suministró esta información. Del mismo modo, se agregó la variable explicativa Facultad a la que ingresa el estudiante, y ésta se clasificó de acuerdo a los programas que conforman cada una de las cinco Facultades académicas con que cuenta la Universidad de los Llanos. Otro cambio sustancial fue excluir a aquellos estudiantes que presentaron información incompleta, provocando que el número de observaciones pasara de 631 a 574. Estos cambios ayudaron a que el modelo se ajustara mejor a los datos analizados.

En cuanto al software para el análisis de la información, se consideró, además del propuesto (*Gretl*), *Minitab 18* como complemento, debido a la ausencia de algunas pruebas de significancia y bondad de ajuste en el primero.

5. SEGUIMIENTO DEL CRONOGRAMA DE ACTIVIDADES A LA FECHA.

ACTIVIDADES	OBJETIVO RELACIONADO	PERIODO DE EJECUCIÓN	DIFICULTADES (si aplica)	PLAN DE CONTINGENCIA (si aplica)
Elaboración de la propuesta y revisión de la bibliografía.	Ninguno	10 de junio del 2018 a 15 de agosto del 2018	Inconvenientes en cuanto a la no claridad de los formatos para elaborar la propuesta de trabajo de grado.	Asesoría y suministro de los formatos por parte de la Oficina de Investigaciones
Procesamiento de los datos, análisis de la información y análisis de los resultados.	General y específicos	10 de agosto del 2018 al 25 de agosto de 2018	Se encontró que mucha de la información que reportan los estudiantes es falsa, asimismo, se identificó que	Verificar cada información de acuerdo a información suministrada por la plataforma Bitácora del Programa de

			muchos de los estudiantes que registraron retiro definitivo, en el periodo 2018-1 aún registraban como vigentes o en retiro temporal. También se encontró problemas con la poca cantidad de información que reportan los estudiantes, dado que hubo que excluir muchas observaciones por ausencia de datos.	Retención Estudiantil Unillanista (PREU).
Elaboración de informe final y artículo.	General y específicos	Del 26 de agosto al 18 de septiembre	Ninguna	Ninguno
Entrega del informe final y socialización	Ninguno	20 de septiembre	Ninguna	Ninguno

6. IMPACTO SOCIAL

TIPO DE IMPACTO	DESCRIPCIÓN DEL IMPACTO	PROYECCIÓN DEL IMPACTO
Científico	Que el documento final sirva como referente para futuras investigaciones y posteriores estudios.	Consultas y citaciones del documento.
Económico	Que los resultados de la investigación sirvan como insumo para implementar planes y/o estrategias enfocadas a ampliar la posibilidad de permanencia y graduación de los estudiantes de la Universidad de los Llanos.	Implementación de planes que promuevan bajas tasas de deserción en la Universidad de los Llanos.

7. DISCUSIÓN Y ANÁLISIS

Los datos fueron procesados utilizando los paquetes de análisis estadístico y econométrico *Gretl* y *Minitab 18*. En la Tabla 1 se muestra el índice de deserción temprana por Facultad. En la Tabla 2 se muestran los coeficientes estimados para las variables independientes, así como el valor Wald. Para la selección de las variables a incluir en el modelo se trabajó con un nivel de significancia ($P < 0,05$). En la Tabla 3 se presentan las relaciones de probabilidad de las variables independientes. En la Tabla 4 se presentan los resultados de las pruebas de bondad de ajuste para validar el modelo, a saber: Estadístico Razón de Verosimilitud (ERV) (Jennings, 1986; Silva y Molina, 2017), la prueba de Hosmer y Lemeshow (1980). Finalmente, en la Tabla 5 se presentan los criterios de sensibilidad y especificidad del modelo (Fernández y Pérez, 2005). Una discusión más completa de estas pruebas y de los supuestos de las mismas, se pueden encontrar en Hausman y McFadden (1984) y Small y Hsiao (1985).

Índice de deserción temprana por Facultad

Tabla 1. Índice de deserción temprana por Facultad

Facultad	Índice de deserción temprana*
Facultad de Ciencias Económicas	30,5%
Facultad de Ciencias Básicas e Ingeniería	50%
Facultad de Ciencias de la Salud	37,50%
Facultad de Ciencias Agropecuarias y Recursos Naturales	25,35%
Facultad de Ciencias Humanas y la Educación	35,90%
Índice de Deserción Temprana de la Universidad	34,71%

* Para el cálculo de este índice se toman aquellos estudiantes que desertaron durante los primeros cuatro semestres sobre el total de matriculados en primer semestre en cada Facultad.

Fuente: Cálculos propios a partir de información suministrada por el Programa de Retención Estudiantil Unillanista

Como se puede apreciar en la Tabla 1, la facultad que mayor tasa de deserción temprana presenta es la Facultad de Ciencias Básicas e Ingeniería, seguido de la Facultad de Ciencias de la Salud, la Facultad que menor tasa de deserción presenta es la Facultad de Ciencias Agropecuarias y Recursos Naturales. Sin embargo, es alarmante el considerar que, durante los primeros cuatro

semestres, según lo indica el índice de deserción temprana de la Universidad se retiran cerca del 35% de los estudiantes que se matriculan a primer semestre.

Estimación del modelo logit

Tabla 2. Influencia de las variables independientes en la probabilidad de desertar

Variables independientes	Coefficiente	Error estándar	Wald
Constante	3,77***	1,06	12,65
ICFES	-0,0755***	0,0181	17,40
Facultad¹			
FCBEI	1,392***	0,312	19,91
FCE	0,246	0,273	0,81
FCHYE	0,37	0,311	1,415
FCS	1,233***	0,37	11,105
Sexo			
Mujer	-0,451**	0,197	5,241
Tipo de plantel			
Publico	-0,817***	0,243	11,31
Repitió años			
Si	0,727***	0,232	9,82
Estudios antes			
Si	-0,422**	0,215	3,852
Situación padres			
Separados	0,486***	0,189	6,612

*** (P<0,001) ** (P<0,05)

Fuente: Elaboración propia

Los resultados de la Tabla 2 dan cuenta de que todas las variables incluidas en el modelo son significativas a un nivel de significancia ($P<0,05$). Además, indican que a medida que aumenta el puntaje de la prueba Saber 11, el ser mujer, el haberse graduado de estudios anteriores y el egresar de una institución educativa de carácter público disminuyen la probabilidad de deserción temprana. Asimismo, el haber reprobado años académicos y el hecho de que los padres del estudiante estén separados aumenta las probabilidades de desertar; del mismo modo, los resultados indican que de todas las facultades en las que se puede matricular el estudiante, las que mayor riesgo tienen de deserción temprana son la de Ciencias Básicas e Ingeniería y la Facultad de Ciencias de la Salud.

Para comprobar que efectivamente los coeficientes estimados tienen alguna significancia estadística, se construyó el denominado estadístico de Wald (W), que sigue una distribución Chi-cuadrado con un grado de libertad y un nivel de significancia ($P < 0,005$), en donde la hipótesis nula es que la variable independiente no explica la dependiente. Por ejemplo, el estadístico de Wald para la variable Sexo es ($W=5,241$), al comparar este valor con el de la tabla de valores Chi-Cuadrado con un grado de libertad y ($P < 0,005$), se puede comprobar que $W > 3,84$, por lo que se rechaza la hipótesis nula y se puede afirmar que la variable independiente incide en la variable dependiente.

Entonces, el modelo estimado es el siguiente:

$$Y' = 3,77 - 0,0755ICFES + 1,392FCBEI + 0,246FCE + 0,370FCHYE + 1,233FCS - 0,451Sexo - 0,817Tipo_plantel + 0,727Repitió_años - 0,422Estudios_antes + 0,486Situación_padres \quad (1)$$

Asimismo,

$$Ln \left[\frac{P(Retiro\ definitivo)}{P(Vigente)} \right] = Y' \quad (2)$$

Por lo tanto,

$$Prob(Retiro\ definitivo) = \frac{exp^{Y'}}{(1 + exp^{Y'})} \quad (3)$$

Relaciones de probabilidad

Las relaciones de probabilidades, también llamadas Odds Ratio (OR), corresponden a un cociente entre dos odds, siendo un odds una forma alternativa de expresar la posibilidad de ocurrencia de un evento de interés o de presencia de una exposición. Una discusión más completa de lo que son las Odds Ratio se encuentra en (Cerdeira *et al.*, 2013; y Stare y Maucourt-Boulch, 2016).

La razón de Odds estimada, por ejemplo, de desertar habiéndose matriculado en la Facultad de Ciencias Básicas e Ingeniería, en comparación con un estudiante que se matricula en la Facultad de Ciencias Agropecuarias y Recursos Naturales es, 4,0246 (Tabla 3). Lo que indica la Odds Ratio estimada es que un alumno que se matricule en la Facultad de Ciencias Básicas e Ingeniería tiene cerca de 4 veces más riesgo de desertar que uno que se matricule en la Facultad de Ciencias Agropecuarias y Recursos Naturales.

En la Tabla 3 se muestran los valores de las Odds Ratio para los distintos grupos de variables; estos valores se muestran para dos niveles; a saber: nivel A y nivel B. El nivel B es el nivel de referencia para el factor. Una OR mayor que 1 indican que el evento es menos probable en el nivel B. Una OR inferior a 1 indica que el evento es más probable en el nivel de referencia, y una OR igual a 1 implica que las dos categorías comparadas son iguales (Aedo *et al.*, 2010).

Tabla 3. Relaciones de probabilidad para las variables independientes

Variable	Unidad de cambio	Odds ratio	IC de 95%
ICFES	1	0,9273	(0,8950. 0,9607)

Nivel A	Nivel B	Odds ratio	IC de 95%
Facultad			
FCBEI	FCAYRN	4,0246	(2,1815. 7,4252)
FCE	FCAYRN	1,2786	(0,7482. 2,1851)
FCHYE	FCAYRN	1,4472	(0,7860. 2,6647)
FCS	FCAYRN	3,43	(1,6621. 7,0785)
FCE	FCBEI	0,3177	(0,1862. 0,5421)
FCHYE	FCBEI	0,3596	(0,1946. 0,6643)
FCS	FCBEI	0,8523	(0,4124. 1,7614)
FCHYE	FCE	1,1318	(0,6746. 1,8988)
FCS	FCE	2,6826	(1,4114. 5,0985)
FCS	FCHYE	2,3701	(1,1577. 4,8521)
Sexo			
Mujer	Hombre	0,6369	(0,4328. 0,9373)
Tipo de plantel			
Publico	Privado	0,442	(0,2745. 0,7115)
Repitió años			
Si	No	2,0688	(1,3141. 3,2569)
Estudios antes			
Si	No	0,6556	(0,4302. 0,9993)
Situación padres			
Separados	Conviven	1,6252	(1,1220. 2,3540)

Fuente: Elaboración propia

Bondad de ajuste del modelo

Tabla 4. Estadísticos de bondad de ajuste del modelo

-2Ln [L(R)]	-2Ln[L(MV)]	ERV	Test de Hosmer-Lemeshow
786,166	704,140	$\chi^2 = 82,026$ P(0,000)	$\chi^2 = 3,31$ (gl=8) P(0,913)

Fuente: Elaboración propia

Para evaluar la significación global del modelo se utiliza el Estadístico Razón de Verosimilitud (ERV), donde L(R) es la función de verosimilitud para el modelo restringido en el que únicamente se considera la constante; y L(MV) es la función de verosimilitud para el modelo formulado.

Este estadístico sigue una distribución Chi-cuadrado con (k-1) grados de libertad, donde K es el número de parámetros incluidos en el modelo formulado y que han sido estimados por máxima verosimilitud, en donde la hipótesis nula es que todos los coeficientes del modelo estimado, excepto la constante son nulos. Esta hipótesis se rechaza si ERV excede el valor crítico. En la Tabla 4 se muestra el valor de ERV y el nivel (P), ante estos resultados, se puede comprobar que efectivamente, la deserción temprana se relaciona con el conjunto de variables estimado.

Otra prueba de bondad de ajuste es la prueba de Hosmer-Lemeshow, que consiste en comparar las frecuencias observadas y esperadas de eventos y no eventos para evaluar hasta qué punto el modelo se ajusta a los datos. La prueba de Hosmer-Lemeshow sigue una distribución Chi-cuadrado con (n-k) grados de libertad y ($P < 0,05$), que tiene como hipótesis nula: El modelo seleccionado ajusta bien los datos (Hosmer y Lemeshow, 1980).

Si el valor p para la prueba de bondad de ajuste es menor que el nivel de significancia elegido, las probabilidades pronosticadas se desvían de las probabilidades observadas de una manera que la distribución binomial no predice. En la Tabla 4 se muestra el valor del test Hosmer-Lemeshow (3,31) con un nivel ($P = 0,913$), ante estos resultados, se acepta el modelo como bueno.

Otra medida de bondad de ajuste que, así como el test de Hosmer-Lemeshow intenta juzgar la precisión con la que el modelo estimado se aproxima a los datos observados, es, la tasa de predicciones correctas. En la Tabla 5 se aprecia la clasificación que hace el modelo seleccionado y se verifican los criterios de especificidad y sensibilidad del modelo.

El punto de corte para la clasificación fue 0,5. Este punto de corte implica que si la probabilidad de desertar es timada por el modelo es superior a 0,5, se pronostica que el estudiante tiene riesgo de desertar durante los primeros cuatro semestres, o que no lo hace en caso contrario, esto es, que para valores menores a 0,5 se pronostica que el estudiante no deserta durante los primeros cuatro semestres.

Tabla 5. Pronósticos del Modelo

Estado	Vigente	Retiro	% de aciertos
Vigente	247	77	76%
Retiro	115	135	54%
Porcentaje global			67%

Fuente: Elaboración propia

En general, se encontró, que el modelo propuesto tiene una tasa de predicciones correctas del 67%, su nivel de especificidad es del 76%, es decir, clasificó correctamente el 76% de los estudiantes que permanecen vigentes, y su nivel de sensibilidad es del 54%, es decir, clasificó correctamente el 54% de los estudiantes que presentaron deserción temprana.

CONCLUSIONES

El modelo estimado evidencia que un buen puntaje en la prueba Saber 11, el ser mujer, no haber reprobado años durante el bachillerato, el haber cursado estudios antes y si los padres conviven, disminuye la probabilidad de deserción temprana; asimismo, el haber egresado de un colegio privado aumenta la probabilidad de deserción. De todas las facultades en las que se puede matricular el estudiante, las que mayor riesgo tienen de deserción temprana son la de Ciencias Básicas e Ingeniería y la Facultad de Ciencias de la Salud.

Asimismo, de acuerdo a las pruebas de bondad de ajuste se considera que el modelo se adapta a los datos observados. Sin embargo, el resultado global de predicción es relativamente bajo, lo que indica que se pueden llegar a incluir otras variables que aporten explicación a la probabilidad de deserción temprana.

Esta investigación corrobora que la tasa de deserción temprana en la Universidad de los Llanos es alta y, sobre todo, logra identificar qué programas presentan los índices más elevados, con estos resultados, se puede esperar que se refuercen y focalicen los planes anti deserción para que se logre disminuir estas cifras.

Las recomendaciones de política educativa que puede aplicar la Universidad de los Llanos se pueden enfocar en realizar pruebas vocacionales para que el estudiante que ingresa acierte en la elección de programa en el que se desea matricular, del mismo modo, es de vital importancia que al detectar a un estudiante en riesgo de deserción se le brinden todas las ayudas disponibles para que este, bien sea, no deserte o tome una mejor decisión. Asimismo, se recomienda una revisión más detallada de la información que suministran los estudiantes cuando realizan la inscripción para que esta pueda ser utilizada con mayor precisión en la detección de posibles desertores.

Se recomienda también, realizar estudios de deserción temprana para cada programa para identificar las causas que afectan a los distintos grupos de estudiantes, todo con el fin de obtener mejores resultados cuando se apliquen, de manera focalizada los distintos planes enfocados a prevenir la deserción. Por ejemplo, se puede considerar realizar refuerzos en áreas como la matemática y física y cursos nivelatorios o de profundización en otras áreas como la química y estadística.

SIGLAS Y ABREVIATURAS

PREU: Programa de Retención Estudiantil Unillanista

REFERENCIAS BIBLIOGRÁFICAS

- Aedo S, Pavlov S, Clavero S. 2010. Riesgo relativo y Odds ratio ¿Qué son y cómo se interpretan? *Revista de obstetricia y ginecología*, 5(1): 51-54.
- Cerda J, Vera C, Rada G. 2013. Odds ratio: aspectos teóricos y prácticos. *Revista médica de Chile*, 41(10): 1329-1335.
- Fernández H, Pérez F. 2005. El modelo logístico: una herramienta estadística para evaluar el riesgo de crédito. *Revista Ingenierías Universidad de Medellín*, 4(6): 55-75
- Hausman JA, McFadden D. 1981. Specification Tests for the Multinomial Logit Model. *Econometrica*, 52(5): 1219-1240.
- Hosmer DW, Lemeshow S. 1980. Goodness of fit tests for the multiple logistic regression model. *Communication in Statistics- Theory and Methods*, 9(10): 1043-1069.
- Jennings DE. 1986. Judging inference adequacy in logistic regression. *Journal of The American Statistical Association*, 81: 471-476
- Silva A, Molina M. 2017. Likelihood ratio (razón de verosimilitud): definición y aplicación en Radiología. *Revista Argentina de Radiología*, 81(3): 204-208.
- Small KA, Hsiao C. 1985. Multinomial Logit Specification Tests. *International Economic Review*, 26(3): 619-627.
- Stare J, Maucourt-Boulch D. 2016. Odds Ratio, Hazard Ratio and Relative Risk. *Metodoloski zvezki*, 13(1): 59-67.

LISTA DE ANEXOS

Anexo 1. Manuscrito de artículo científico sometido para publicación

**Análisis de la deserción temprana y sus factores explicativos en la Universidad de los Llanos
con datos de ingreso 2015-2**

**Analysis of the early desertion and his explanatory factors in the University of the Llanos
with information of revenue 2015-2**

**Análise do abandono precoce e seus factores explicativos na Universidade de Llanos com
detalhes de login de 2015-2**

Cristian A Arismendy-Fuentes*¹

Nathalia L Morales-Parrado²

1-2. Estudiante de Economía, Universidad de los Llanos, Villavicencio, Colombia.

* cristian.arismendy@unillanos.edu.co

RESUMEN

Esta investigación tiene como principal propósito analizar la probabilidad de deserción temprana de los estudiantes de la Universidad de los Llanos mediante un modelo de regresión logística. Además, identificar cuáles son las variables de mayor incidencia en la deserción temprana. Para la construcción del modelo se tomó como muestra la información de 574 estudiantes que ingresaron en la cohorte 2015-2 y que para el periodo 2018-1 eran registrados como vigentes o reportaban como último periodo matriculado cualquiera de los cuatro primeros semestres. Los resultados dan cuenta de que un buen puntaje en la prueba Saber 11, el ser mujer, no haber reprobado años durante el bachillerato, el haber cursado estudios antes y si los padres conviven, disminuye la probabilidad de deserción temprana; asimismo, el haber egresado de un colegio privado aumenta la probabilidad de deserción. De todas las facultades en las que se puede matricular el estudiante, las que mayor riesgo tienen de deserción temprana son la de Ciencias Básicas e Ingeniería y la Facultad de Ciencias de la Salud.

Palabras clave: *deserción temprana, modelo logit, Universidad de los Llanos*

ABSTRACT

This research has as main purpose to analyze the likelihood of early dropout of students from the University of the Llanos using a logistic regression model. In addition, identify which variables are most prevalent early dropout. To build the model was sampled information from 574 students who entered the cohort 2015-2 and 2018-1 for the period were recorded as current or last period reported as enrolled any of the first four semesters. The results show that a good test score Saber 11 being female, not failed years in high school, having studied before and if the parents live together, decreases the likelihood of early dropouts; Likewise, Graduates have a private school increases the likelihood of dropping out. Of all the faculties which can enroll the student, which have increased risk of early dropouts are the Basic Sciences and Engineering and the Faculty of Health Sciences.

Keywords: *early dropout, logit model, Universidad de los Llanos*

RESUMO

Esta pesquisa tem como objetivos analisar a probabilidade de abandono precoce de estudantes da Universidade do Llanos usando um modelo de regressão logística. Além disso, identificar quais variáveis são mais prevalentes abandono precoce. Para construir o modelo foi amostrado informações de 574 alunos que ingressaram na coorte 2015-2 e 2018-1 para o período foram registrados como período atual ou a última relatado como inscrito qualquer um dos quatro primeiros semestres. Os resultados mostram que uma boa pontuação no teste Sabre 11 sendo do sexo feminino, anos, não falhou na escola, tendo estudado antes e se os pais vivem juntos, diminui a probabilidade de abandono escolar precoce; Além disso, Os graduados têm uma escola particular aumenta a probabilidade de abandonar a escola. De todas as faculdades que pode se inscrever o aluno, que têm maior risco de abandono escolar precoce são as Ciências Básicas e Engenharia e da Faculdade de Ciências da Saúde.

Palavras chave: *abandono precoce, modelo logit, Universidad de los Llanos*

INTRODUCCIÓN

La educación superior en Colombia enfrenta importantes retos, dentro de los que se destacan la ampliación de la cobertura, el mejoramiento de la calidad de las instituciones que ofrecen este nivel de enseñanza (Melo-Becerra, *et al.*, 2017), la disminución de la desigualdad en el acceso entre grupos sociales de diferentes ingresos y la deserción (Jaramillo, 2010). El problema de la deserción universitaria cobra importancia debido a que esta situación dilata el tiempo de formación de capital humano de calidad en el país. Autores como Becker (1964) y Hanushek (1986) consideran que una mejor educación se relaciona con mayores dotaciones de capital humano de un país que supondrá una mayor predisposición para igualar las condiciones de los miembros de una sociedad.

Las altas tasas de deserción dejan en evidencia la poca capacidad de retención de los centros educativos al no poder mantener en sus aulas a todos los estudiantes que se matriculan. Según datos del Sistema para la Prevención de la Deserción de la Educación Superior (SPADIES), en el 2015, la tasa de deserción nacional por cohorte, a nivel de estudios universitarios era del 41,6%, lo que indica, que tan solo el 58,4% de los estudiantes que se matriculan alcanzan décimo semestre.

Para la Universidad de los Llanos la situación no es distinta, según el boletín estadístico de la Universidad de los Llanos (2018), la tasa de deserción por periodo para la cohorte 2017-1 es del 8%, este valor representa el porcentaje de estudiantes que estando matriculados son clasificados como desertores un año después si no renuevan matrícula, del mismo modo, el boletín indica que la deserción ponderada acumulada por cohorte para los periodos 1998-1 a 2015-2 es del 17,90% en el primer semestre, llegando al 45,93% en décimo semestre, esta información muestra el promedio ponderado de deserción por cohorte sin ligar la cifra a un periodo de ingreso, sino a nivel institucional. Por lo tanto, indica que en promedio el 17,90% de los estudiantes que se matricularon durante estos periodos decidieron desertar habiendo realizado máximo un semestre académico. Y, asimismo, al llegar a decimo semestre en promedio solo el 54,07% de los estudiantes que se matricularon durante estos periodos en primer semestre siguen vigentes o lograron graduarse. Del mismo modo, para el periodo de estudio (2015-2, 2018-1) se encontró que durante los cuatro primeros semestres se retiran cerca del 35% de los estudiantes matriculados (Tabla 1.).

Tabla 1. Deserción por periodo acumulada

	Periodo					
	2015-2	2016-1	2016-2	2017-1	2017-2	2018-1
Deserción por periodo	68	37	86	61	24	9
Deserción por periodo acumulada	68	105	191	252	276	285
Índice de deserción por periodo	9,37%	5,10%	11,85%	8,40%	3,31%	1,24%
Índice de deserción por periodo acumulada	9,37%	14,46%	26,31%	34,71%*	38,02%	39,26%
Total matriculados	726	658	621	556	495	424**

*Índice de deserción temprana

**Incluye el total de estudiantes que al periodo 2018-1 se encuentran vigentes, en retiro temporal, o han aplazado matrícula, y excluye a aquellos que ya se graduaron (Ver Tabla 2.)

Fuente: Cálculos propios a partir de información suministrada por el Programa de Retención Estudiantil Unillanista

Aunque la deserción de estudios universitarios ha sido ampliamente estudiada, a esta se le ha vinculado con distintos factores, desde aspectos socioeconómicos, capacidades individuales (Tinto, 1975), factores psicosociales, ambientales, de socialización, y actitudinales (Bean, 1981), organizacionales y de interacción con el centro de educación superior (Braxton, *et al.*, 1997), Si bien se ha avanzado en estrategias enfocadas a prevenir, evitar y reducir la deserción; las causas que la producen son múltiples, y en ocasiones, determinarlas no es tarea sencilla, llevando consigo que la prevención del problema sea poco efectiva.

El objetivo de este trabajo es estimar, a partir de un modelo de regresión logística la probabilidad de que un estudiante que se matricule en la universidad de los llanos deserte durante los primeros cuatro semestres; y, a su vez, identificar los factores que inciden en la deserción temprana. El modelo es estimado con diversas características, incluyendo características generales de los estudiantes, aspectos socioeconómicos, características familiares o de entorno, y características académicas.

Estado del arte

Abundante es la literatura dedicada al fenómeno de la deserción estudiantil y múltiples han sido los estudios que se han realizado para identificar sus causas, éstos han sido analizados e interpretados desde distintas disciplinas.

Una de las definiciones más aceptadas de lo que es la deserción la ofrece Tinto (1975), quien la define como el proceso de abandono voluntario o forzoso de la carrera en que se matricula un estudiante, por la influencia positiva o negativa de circunstancias internas o externas al alumno. Asimismo, Himmel (2002) se refiere a la deserción como el abandono prematuro de un programa de estudios antes de alcanzar el título o grado, y considera un tiempo suficientemente largo como para descartar la posibilidad de que el estudiante se reincorpore.

En relación a las causas que producen la deserción estudiantil, múltiples estudios han demostrado que este fenómeno no se puede atribuir a un solo factor, en cambio, se consideran un conjunto de factores que se relacionan entre sí. Por ejemplo, Tinto (1975) encuentra que las variables que explican la deserción están relacionadas con las capacidades desarrolladas por el estudiante, previo a su ingreso a la educación superior, del mismo modo Bean (1981) señala que la deserción no es atribuible exclusivamente variables académicas, sino que puede estar explicada por factores psicosociales, ambientales, de socialización, y actitudinales. En esta línea, Jadue (2002) analiza desde un punto de vista clínico algunos factores psicológicos y sociales que aumentan el riesgo estudiantil de bajo rendimiento y provocan deserción en la escuela. Asimismo, Acevedo, et al., (2015) consideran que la mayoría de los estudiantes deciden interrumpir los estudios por causa de su situación económica. Del mismo modo, Larroucau (2015), constata que la deserción universitaria también es un problema de financiamiento, encontrando que los estudiantes de mayores ingresos y aquellos que perciben una ayuda estudiantil presentan una probabilidad mucho menor de desertar. Por otro lado, Patiño y Cardona (2013) encontraron que los exámenes de ingreso, la orientación vocacional, la economía y los problemas personales presentan mayor relevancia a la hora de desertar de algún programa académico. Otros estudios como los de Giovagnoli (2002) y Porto y Di Gresia (2004) consideran que el éxito académico depende de la educación de los padres, el tipo de escuela secundaria y, aseguran que iniciar la carrera y estar trabajando hace que el estudiante tenga más riesgos de abandonar. De igual manera, estos autores concluyen en que la deserción es mayor en los primeros semestres de la carrera universitaria.

En esta línea, Patiño y Cardona (2013) encontraron que a nivel de latinoamericano los exámenes de ingreso, la orientación vocacional, la economía y los problemas personales presentan mayor relevancia. Además, concluyeron que la falencia para rebatir la deserción yace en la ausencia de políticas efectivas que frenen la deserción y en la falta de apoyo a las políticas educativas regionales para integrar a la universidad, la sociedad y el Estado. Asimismo, Canales y De los Ríos (2018) identificaron que la deserción temporal se explica principalmente por razones vocacionales, socioculturales y motivacionales, mientras que la permanente, por razones socioeconómicas.

Para el caso Colombiano, la mayoría de los estudios se centran en casos concretos de distintas Universidades, en donde se pretenden identificar los factores específicos que causan la deserción. En esta línea, Sánchez et al., (2002) en un trabajo que buscó identificar los factores determinantes en el acceso y la permanencia en las universidades públicas colombianas encontraron que el acceso a las universidades públicas está determinado, en principio, por el resultado en las pruebas de Estado. De otro lado, el estudio muestra que la existencia de programas de "acción afirmativa" durante la carrera aumenta en forma positiva y significativa la probabilidad de permanencia de los estudiantes en general, aunque no aumenta la permanencia de los estudiantes más pobres.

Más recientemente, Olave et al., (2013) ponen en relación la deserción universitaria con la alfabetización académica, entendida como el nivel de desarrollo de la lectura y la escritura. Por otro lado, Acevedo y Bayona (2011) identificaron factores motivacionales y sociodemográficos de estudiantes desertores de la facultad de psicología de la universidad Pontificia Bolivariana en Bucaramanga, encontrando que tanto la parte familiar como la personal, se relaciona con los desertores en general para que estos dejaran de estudiar; mientras que en los desertores de universidad, los problemas con sus docentes y con la institución hicieron que tomaran la decisión de cambiar de Universidad. En lo sociodemográfico, estos autores encontraron que la deserción se halla relacionada con el nivel educativo de los padres, un historial de alta movilidad y falta de orientación vocacional. Y concluyen afirmando que la deserción no tiene que ver en su totalidad con el rendimiento académico.

Para el caso de la Universidad de los Llanos, la literatura al respecto es extremadamente limitada, tan solo los trabajos de (Malagón et al., 2006 y Malagón et al., 2007) en los cuales se procesó información correspondiente de las cohortes 1998-2004, y así mismo se determinaron índices de deserción por programas académicos como análisis longitudinal y transversal; posteriormente,

además, se analizaron los factores generales que inciden en la deserción para cuantificar el impacto de tales causas sobre la decisión de abandonar los estudios, dados por factores socioeconómicos, académicos, personales e institucionales. Los autores encontraron que el 44.0% de los estudiantes que ingresaron a la Universidad de los Llanos en las diferentes cohortes no terminaron sus estudios, siendo las Ingenierías de Sistemas y Electrónica las más afectadas por el abandono de sus estudiantes con un 56.0% y 55.0% respectivamente, el programa que presentó menor Índice de Deserción fue la Licenciatura en Educación Física y Deportes (33.0%). Del mismo modo encontraron que las causas que generaron mayor deserción en la Universidad de los Llanos durante el periodo estudiado son las académicas con un 48.0% frente al 26.0% que desertaron por causas socioeconómicas que incluyen el abandono por la baja condición económica del estudiante.

Estos trabajos ofrecen una mirada general de cómo se ha producido históricamente la deserción en la Universidad de los Llanos, sin embargo, carecen de un análisis estadístico profundo y no ofrecen una propuesta rigurosa de cómo tratar los resultados obtenidos. Además, presentan de manera descriptiva las variables concernientes a los estudiantes de las cohortes de estudio, pero no logran relacionarlas para identificar cómo afectan en conjunto la decisión de desertar de los estudiantes. Por lo tanto, esta información no es suficiente para determinar cuáles son las causas que realmente importan y el grado de incidencia que éstas tienen en la deserción estudiantil.

MATERIALES Y MÉTODOS

Diseño

Con base en la información que recopila cada semestre la Oficina de Admisiones, Registro y Control, el Programa de Retención Estudiantil Unillanista y la Oficina de Planeación de la Universidad de los Llanos a los estudiantes que se matriculan en la Universidad, se seleccionó como población objeto de estudio aquellos estudiantes que efectivamente se matricularon en la cohorte 2015-2; en la selección de la información no se tomó en cuenta a los estudiantes del programa de Licenciatura en Producción Agropecuaria y los de la tecnología en Gestión de Empresas Pecuarias, debido a que a la fecha no se volvieron a abrir inscripciones para estos programas. En la Tabla 2 se presenta el estado de los estudiantes al periodo 2018-1, es importante

aclarar, que de los estudiantes que se encuentran en retiro definitivo, 252 lo hicieron durante los primeros cuatro semestres. En la selección de la muestra se tuvo en cuenta solo a aquellos estudiantes quienes para la cohorte 2018-1 hacen parte de estas dos categorías:

1. Estudiante vigente: estudiante que registra como último periodo matriculado 2018-1.
2. Estudiante desertor: aquel que radicó su retiro definitivo de la universidad sin discriminación de causa, o aquel que estando matriculado en el periodo t-2 no registra matrícula en el periodo t-1 ni en el periodo t.

Tabla 2. Estado de los estudiantes a 2018-1

Estado	Frecuencia [*]	Proporción
Vigente	401	55,23%
Retiro Definitivo	285	39,26%
Retiro Temporal	14	1,93%
Aplazado	9	1,24%
Graduado	17	2,34%
Total	726	100%

**No se incluyeron Licenciatura en Producción Agropecuaria y la Tecnología en Gestión de Empresas Agropecuarias*

Fuente: Fuente: Cálculos propios a partir de información suministrada por el Programa de Retención Estudiantil Unillanista.

Esta diferenciación de categorías excluye a aquellos estudiantes quienes a la fecha de selección de la información registran como graduados, en matrícula aplazada o en estado de retiro temporal. Además, también se excluyen aquellos estudiantes que realizaron retiro definitivo en quinto o sexto semestre, debido a que sólo se pretende estimar las probabilidades de deserción temprana; es decir, aquellos estudiantes que registraron como último periodo matriculado cualquiera de los cuatro primeros semestres de carrera. Quedando finalmente una muestra de 574 estudiantes que cumplían con las características antes descritas. En la Tabla 3 se puede apreciar la descripción de las variables dicotómicas.

Tabla 3. Descripción de las variables dicotómicas.

Variable	Frecuencia	Porcentaje	Variable	Frecuencia	Porcentaje
Estado			Repitió años		
Retiro definitivo	250	43,55	No	461	80,31
Vigente	324	56,45	Si	113	19,69
Facultad			Situación_padres		
FCAYRN	106	18,47	Conviven	226	39,37
FCBEI	105	18,29	Separados	348	60,63
FCE	209	36,41	Tipo de plantel		
FCHYE	98	17,07	Privado	97	16,9
FCS	56	9,76	Publico	477	83,1
Sexo			Estudió antes		
Hombre	288	50,17	No	414	72,13
Mujer	286	49,83	Si	160	27,87

Fuente: Elaboración propia

Modelos de Respuesta Cualitativa

En economía, muchos comportamientos son cualitativos, por lo que para explicarlos deberá ampliarse el modelo de regresión clásico, de manera que incluya también variables cualitativas (Lassibille y Navarro, 1981). Dentro de la literatura estadística gozan de reconocimiento unos modelos que permiten operar con estas variables de respuesta cualitativa, entre los que se destacan, el modelo probit y el modelo logit o también llamado modelo de regresión logística (Amemiya, 1975; Núñez, 1990). Cada uno con características distintas y aplicables según sea el requerimiento.

La regresión logística

Según Hosmer y Lemeshow (2000), lo que distingue la regresión logística de los modelos de regresión lineal, es fundamentalmente, que la variable de respuesta es dicotómica o binaria, es decir, toma uno de dos valores posibles. Al igual que otras técnicas estadísticas multivariadas, el modelo de regresión logística da la posibilidad de evaluar la influencia de cada una de las variables independientes sobre la variable dependiente o de respuesta y controlar el efecto del

resto (Alderete, 2006). Los objetivos del modelo de regresión logística según Salas (1996) son tres:

1. Determinar la existencia o ausencia de relación entre una o más variables independientes (X_i) y una variable dependiente dicotómica (Y).
2. Medir el signo de dicha relación en caso de que exista.
3. Estimar o predecir la probabilidad de que se produzca el suceso o acontecimiento definido como ($Y=1$) en función de los valores que adoptan las variables independientes.

El modelo más sencillo para determinar qué características influyen en la permanencia estudiantil, es aquel que incluye una sola variable independiente X :

$$Y = \alpha + \beta X + \mu \quad (1)$$

Donde α es el término constante o independiente, β es el coeficiente de regresión asociado a la variable independiente X , y μ es el término de perturbación aleatoria o error del modelo.

Pero como en este caso, Y es una variable dicotoma que toma valores de 1 (en caso de desertar durante los primeros cuatro semestres) o 0 (en caso de permanecer vigente), entonces para cada observación la perturbación μ debe ser una variable aleatoria que solo puede tomar los valores $[1 - (\alpha + \beta X)]$ y $[-(\alpha + \beta X)]$, respectivamente. Además, para que $E(\mu) = 0$, las probabilidades con que μ debe tomar estos dos valores han de ser $(\alpha + \beta X)$ y $[1 - (\alpha + \beta X)]$, respectivamente.

Por lo tanto, se estaría interesado en estimar el siguiente modelo:

$$P(Y = 1) = \alpha + \beta X \quad (2)$$

Donde P es la probabilidad estimada de que un estudiante seleccionado al azar deserte durante los primeros cuatro semestres. Sin embargo, el modelo $P(Y = 1) = \alpha + \beta X$ no cumple con la condición $0 \leq E(Y_i|X_i) \leq 1$. Por lo que se estaría tratando de un Modelo Lineal de Probabilidad (MLP); el incumplimiento de la condición se debe a que si se realiza la estimación del modelo por el método de Mínimos Cuadrados Ordinarios (MCO), éste no tiene en cuenta la restricción $0 \leq E(Y_i) \leq 1$ ante esto, se podrían encontrar valores $(\hat{Y}_i > 1)$ y $(\hat{Y}_i < 0)$.

Ante esta situación, lo que se busca es ajustar el modelo para que efectivamente garantice que las probabilidades estimadas se encuentren con seguridad entre 0 y 1. Así:

$$P(Y = 1) = \frac{\exp^{(\alpha + \beta X)}}{1 + \exp^{(\alpha + \beta X)}} = \frac{1}{1 + \exp^{-(\alpha + \beta X)}} \quad (3)$$

A esta expresión se le conoce como *Función de Regresión Logística* (FRL), donde $\exp(.)$ representa la función exponencial. La (FRL) también se puede expresar de la siguiente forma:

$$\frac{P(Y = 1)}{1 - P(Y = 1)} = \exp^{(\alpha + \beta X)} \quad (4)$$

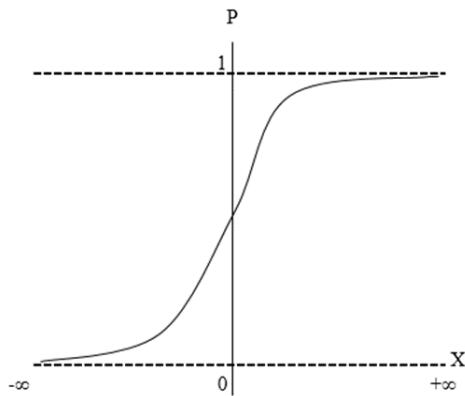
Sin embargo, como la (FRL) se expresa de manera exponencial, esta puede transformarse en una función lineal tomando logaritmos naturales, quedando así:

$$\ln \left[\frac{P(Y = 1)}{P(Y = 0)} \right] = \alpha + \beta X \quad (5)$$

Ahora $[P(Y=1))/P(Y=0)]$ es lo que se conoce como *Odds* o simplemente la razón de las probabilidades en favor de que un estudiante deserte durante los primeros cuatro semestres respecto a la probabilidad de que permanezca vigente.

Así mismo la función $\ln[(P(Y=1))/(P(Y=0))]$ vendría siendo el logaritmo de las *Odds Ratio* y esta se conoce como *logit* (L). Los posibles valores que tome (L) pueden oscilar entre $(-\infty)$ si $P(Y=1) = 0$, y $(+\infty)$, cuando $P(Y=1) = 1$. Del mismo modo, aunque (L) es lineal en X, las probabilidades en sí mismas no lo son, a diferencia del MLP, en donde las probabilidades aumentan linealmente conforme aumenta X, por lo tanto, la probabilidad estimada $P(Y=1)$ es una función curvilínea en forma de “S”, lo que permite que las estimaciones de probabilidad que se predicen se mantengan en el rango de valores comprendidos entre 0 y 1 como se muestra en la Figura 1.

Figura 1. **Función de la Distribución Logística**



Fuente: Elaboración Propia.

Modelo logístico multivariable

Como se mencionó en la anterior sección, el objetivo básico de la regresión logística es pretender explicar un fenómeno que se puede clasificar de forma dicotómica [$Y=(1,0)$], a partir de un conjunto de variables independientes (X_i) que se consideran relevantes dentro de lo que se supone es la causa del fenómeno estudiado; en este caso, permanecer o desertar de la Universidad de los Llanos.

La introducción de dos o más variables independientes ($X_1, X_2, X_3, X_4, \dots, X_n$) en el modelo de regresión logística, constituye el denominado *Modelo Logístico Multivariable* (MLM):

$$\ln \left[\frac{P(Y = 1)}{P(Y = 0)} \right] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n \quad (6)$$

Este será el modelo que se usará para determinar las probabilidades de deserción de un estudiante durante los primeros cuatro semestres [$P(Y=1)$] y la probabilidad de permanecer vigente [$P(Y=0)$].

Análisis estadístico

La variable que se quiere explicar es de carácter dicotómico o Dummy, y se define así:

($Y=1$) si el estudiante desertó durante los primeros cuatro semestres de carrera, o registra como último periodo matriculado cualquiera de los primeros cuatro semestres.

($Y=0$) si el estudiante se encuentra vigente en el periodo 2018.1.

Con base en la caracterización que se realiza a los estudiantes que ingresan a primer semestre, se definieron las variables explicativas (X_i) que serán tenidas en cuenta para la elaboración del modelo de regresión logística.

A continuación, se especifica cada una de ellas:

Sexo: Variable cualitativa, que toma el valor de 1 si es Mujer y 0 si es Hombre.

Facultad: Variable cualitativa que se refiere a la facultad académica en la que se matricula el estudiante que ingresa a la universidad de los Llanos. A saber: Facultad de Ciencias Económicas (FCE), Facultad de Ciencias Básicas e Ingeniería (FCBEI), Facultad de Ciencias de la Salud (FCS), Facultad de Ciencias Agropecuarias y Recursos Naturales (FCAyRN) y Facultad de Ciencias Humanas y Educación (FCHyE).

Tipo de plantel: Variable cualitativa que distingue si el estudiante terminó sus estudios secundarios en una institución educativa de carácter público (1) o privada (0).

Repitió años: Variable cualitativa que toma el valor de 1 si el estudiante repitió algún año académico y 0 en caso contrario.

Puntaje del ICFES: Variable cuantitativa que relaciona el puntaje de la prueba saber 11 como requisito para el ingreso a la Universidad de los Llanos.

Estudios antes: Variable cualitativa que distingue si el estudiante realizó estudios antes y logró graduarse (1) o no los realizó (0).

Situación de los padres: Variable cualitativa que distingue si los padres del estudiante están separados (1) o conviven (0).

RESULTADOS Y DISCUSIÓN

Los datos fueron procesados utilizando los paquetes de análisis estadístico y econométrico *Gretl* y *Minitab 18*. En la Tabla 4 se muestra el estado de los estudiantes que ingresaron en la cohorte 2015-2 al periodo 2018-1, en la Tabla 5 se muestra en índice de deserción temprana por Facultad, y en la tabla 6 se muestra el índice de deserción temprana para cada programa que ofrece la Universidad de los Llanos. En la Tabla 7 se muestran los coeficientes estimados para las variables independientes, así como el valor Wald. Para la selección de las variables a incluir en el modelo se trabajó con un nivel de significancia ($P < 0,05$). En la Tabla 8 se presentan las relaciones de probabilidad de las variables independientes. En la Tabla 9 se presentan los resultados de las pruebas de bondad de ajuste para validar el modelo, a saber: Estadístico Razón de Verosimilitud (ERV) (Jennings, 1986; Silva y Molina, 2017), la prueba de Hosmer y Lemeshow (1980). Finalmente, en la Tabla 11 se presentan los criterios de sensibilidad y especificidad del modelo (Fernández y Pérez, 2005). Una discusión más completa de estas pruebas y de los supuestos de las mismas, se pueden encontrar en Hausman y McFadden (1984) y Small y Hsiao (1985).

Índice de deserción temprana por Facultad y programa

Tabla 4. Estado de los estudiantes por programa

Programa	Estado						Total
	Vigente	Retiro temprano	Retiro tardío	Retiro Temporal	Aplazado	Graduado	
Administración de Empresas (horario diurno)	29	15	-	1	-	-	45
Administración de Empresas (horario nocturno)	24	13	4	-	-	-	41
Biología	16	17	4	-	-	-	37
Contaduría Pública (horario diurno)	28	10	2	-	1	-	41
Contaduría Pública (horario nocturno)	30	13	2	2	2	-	49
Economía	25	17	2	2	-	-	46
Enfermería	26	13	4	1	2	-	46
Ingeniería Agroindustrial	28	10	3	-	1	-	42
Ingeniería Agronómica	28	17	1	-	-	-	46
Ingeniería de Sistemas	16	27	4	-	-	-	47
Ingeniería Electrónica	18	20	4	-	1	1	44
Licenciatura en Educación Física	25	17	-	2	-	-	44
Licenciatura en Matemáticas y Física	11	16	-	1	-	-	28
Licenciatura en Pedagogía Infantil	33	9	1	2	-	-	45
Medicina Veterinaria y Zootecnia	36	7	2	1	-	-	46
Mercadeo	22	11	-	1	2	1	37
Tecnología en Regencia de Farmacia	6	20	-	1	-	15	42
Total	401	252	33	14	9	17	726

Fuente: Cálculos propios a partir de información suministrada por el Programa de Retención Estudiantil Unillanista.

Tabla 5. Índice de deserción temprana por Facultad

Facultad	Índice de deserción temprana *****
Facultad de Ciencias Económicas *	30,5%
Facultad de Ciencias Básicas e Ingeniería **	50%
Facultad de Ciencias de la Salud ***	37,50%
Facultad de Ciencias Agropecuarias y Recursos Naturales ****	25,35%
Facultad de Ciencias Humanas y la Educación *****	35,90%
Índice de Deserción Temprana de la Universidad	34,71%

*Economía, Mercado, Administración de Empresas (Diurno-Nocturno), Contaduría Pública (Diurno-Nocturno)

**Biología, Ingeniería de Sistemas, Ingeniería Electrónica

***Enfermería, Tecnología en Regencia de Farmacia

****Ingeniería Agroindustrial, Ingeniería Agronómica, Medicina Veterinaria y Zootecnia

*****Licenciatura en Educación Física y Deportes, Licenciatura en Matemáticas y Física, Licenciatura en Pedagogía Infantil

***** Para el cálculo de este índice se toman aquellos estudiantes que desertaron durante los primeros cuatro semestres sobre el total de matriculados en primer semestre en cada Facultad.

Fuente: Cálculos propios a partir de información suministrada por el Programa de Retención Estudiantil Unillanista.

Como se puede apreciar en la Tabla 5, la facultad que mayor tasa de deserción temprana presenta es la Facultad de Ciencias Básicas e Ingeniería, seguido de la Facultad de Ciencias de la Salud, la Facultad que menor tasa de deserción presenta es la Facultad de Ciencias Agropecuarias y Recursos Naturales.

Al revisar el índice de deserción temprana por programa, se puede evidenciar que Biología, Ingeniería de Sistemas, Ingeniería Electrónica, Licenciatura en Matemáticas y Física y la Tecnología en Regencia en Farmacia presentan tasas por encima del 45%, es decir, un poco más del 45% de los estudiantes que se matriculan en esos programas deciden desertar durante los primeros cuatro semestre de carrera (Tabla 6).

Tabla 6. Índice de deserción por programa

Programa	Índice de deserción temprana*
Administración de Empresas (horario diurno)	33,33%
Administración de Empresas (horario nocturno)	31,71%
Biología	45,95%
Contaduría Pública (horario diurno)	24,39%
Contaduría Pública (horario nocturno)	26,53%
Economía	36,96%
Enfermería	28,26%
Ingeniería Agroindustrial	23,81%
Ingeniería Agronómica	36,96%
Ingeniería de Sistemas	57,45%
Ingeniería Electrónica	45,45%
Licenciatura en Educación Física	38,64%
Licenciatura en Matemáticas y Física	57,14%
Licenciatura en Pedagogía Infantil	20,00%
Medicina Veterinaria y Zootecnia	15,22%
Mercadeo	29,73%
Tecnología en Regencia de Farmacia	47,62%
Índice de deserción Temprana de la Universidad	34,71%

*Para el cálculo de este índice se toman aquellos estudiantes que desertaron durante los primeros cuatro semestres sobre el total de matriculados en primer semestre en cada Programa

Fuente: Cálculos propios a partir de información suministrada por el Programa de Retención Estudiantil Unillanista

Es importante tener en cuenta estos índices para focalizar planes enfocados a prevenir la deserción temprana en los distintos programas que ofrece la Universidad de los Llanos, debido a que las causas que inciden en la deserción no son equivalente para todos los programas. Por ejemplo, se puede considerar realizar refuerzos en áreas como la matemática y física y cursos nivelatorios o de profundización en otras áreas como la química y estadística.

Estimación del modelo logit

Tabla 7. Influencia de las variables independientes en la probabilidad de desertar

Variables independientes	Coeficiente	Error estándar	Wald
Constante	3,77***	1,06	12,65
ICFES	-0,0755***	0,0181	17,40
Facultad¹			
FCBEI	1,392***	0,312	19,91
FCE	0,246	0,273	0,81
FCHYE	0,37	0,311	1,415
FCS	1,233***	0,37	11,105
Sexo			
Mujer	-0,451**	0,197	5,241
Tipo de plantel			
Publico	-0,817***	0,243	11,31
Repitió años			
Si	0,727***	0,232	9,82
Estudios antes			
Si	-0,422**	0,215	3,852
Situación padres			
Separados	0,486***	0,189	6,612

*** (P<0,001) ** (P<0,05)

Fuente: Elaboración propia

Los resultados de la Tabla 7 dan cuenta de que todas las variables incluidas en el modelo son significativas a un nivel de significancia ($P<0,05$). Además, indican que a medida que aumenta el puntaje de la prueba Saber 11, el ser mujer, el haberse graduado de estudios anteriores y el egresar de una institución educativa de carácter público disminuyen la probabilidad de deserción temprana. Asimismo, el haber reprobado años académicos y el hecho de que los padres del estudiante estén separados aumenta las probabilidades de desertar; del mismo modo, los resultados indica que de todas las facultades en las que se puede matricular el estudiante, las que mayor riesgo tienen de deserción temprana son la de Ciencias Básicas e Ingeniería y la Facultad de Ciencias de la Salud.

Para comprobar que efectivamente los coeficientes estimados tienen alguna significancia estadística, se construyó el denominado estadístico de Wald (W), que sigue una distribución Chi-cuadrado con un grado de libertad y un nivel de significancia ($P<0,005$), en donde la hipótesis

nula es que la variable independiente no explica la dependiente. Por ejemplo, el estadístico de Wald para la variable Sexo es ($W=5,241$), al comparar este valor con el de la tabla de valores Chi-Cuadrado con un grado de libertad y ($P<0,005$), se puede comprobar que $W > 3,84$, por lo que se rechaza la hipótesis nula y se puede afirmar que la variable independiente incide en la variable dependiente.

Entonces, el modelo estimado es el siguiente:

$$Y' = 3,77 - 0,0755ICFES + 1,392FCBEI + 0,246FCE + 0,370FCHYE + 1,233FCS - 0,451Sexo - 0,817Tipo_plantel + 0,727Repitió_años - 0,422Estudios_antes + 0,486Situación_padres \quad (7)$$

Asimismo,

$$Ln \left[\frac{P(Retiro\ definitivo)}{P(Vigente)} \right] = Y' \quad (8)$$

Por lo tanto,

$$Prob(Retiro\ definitivo) = \frac{exp^{Y'}}{(1 + exp^{Y'})} \quad (9)$$

Relaciones de probabilidad

Las relaciones de probabilidades, también llamadas Odds Ratio (OR), corresponden a un cociente entre dos odds, siendo un odds una forma alternativa de expresar la posibilidad de ocurrencia de un evento de interés o de presencia de una exposición. Una discusión más completa de lo que son las Odds Ratio se encuentra en (Cerdeira *et al.*, 2013; y Stare y Maucourt-Boulch, 2016).

La razón de Odds estimada, por ejemplo, de desertar habiéndose matriculado en la Facultad de Ciencias Básicas e Ingeniería, en comparación con un estudiante que se matricula en la Facultad de Ciencias Agropecuarias y Recursos Naturales, se estima de la siguiente forma:

$$Odds\ Ratio = \frac{Odds(L1)}{Odds(L0)} \quad (10)$$

Donde,

$$\begin{aligned}L &= \ln(Odds) = \alpha + \beta_i X_i \\ Odds &= \exp^L\end{aligned}\tag{11}$$

Consecuentemente, la Odds estimada de que un alumno de la Facultad de Ciencias Básicas e Ingeniería deserte es:

$$L1 = \ln(Odds) = 3,77 + 1,392(1) = 5,162$$

Del mismo modo, la Odds estimada de que un estudiante de la Facultad de Ciencias Agropecuarias y Recursos Naturales deserte es:

$$L0 = \ln(Odds) = 3,77 + 1,392(0) = 3,77$$

Por lo tanto, la Odds Ratio calculada sería:

$$Odds\ Ratio = \frac{\exp^{5,162}}{\exp^{3,77}} = 4,024$$

Lo que indica la Odds Ratio estimada es que un alumno que se matricule en la Facultad de Ciencias Básicas e Ingeniería tiene cerca de 4 veces más riesgo de desertar que uno que se matricule en la Facultad de Ciencias Agropecuarias y Recursos Naturales.

En la Tabla 8 se muestran los valores de las Odds Ratio para los distintos grupos de variables; estos valores se muestran para dos niveles; a saber: nivel A y nivel B. El nivel B es el nivel de referencia para el factor. Una OR mayor que 1 indican que el evento es menos probable en el nivel B. Una OR inferior a 1 indica que el evento es más probable en el nivel de referencia, y una OR igual a 1 implica que las dos categorías comparadas son iguales (Aedo *et al.*, 2010).

Tabla 8. Relaciones de probabilidad para las variables independientes

Variable	Unidad de cambio	Odds ratio	IC de 95%
ICFES	1	0,9273	(0,8950. 0,9607)

Nivel A	Nivel B	Odds ratio	IC de 95%
Facultad			
FCBEI	FCAYRN	4,0246	(2,1815. 7,4252)
FCE	FCAYRN	1,2786	(0,7482. 2,1851)
FCHYE	FCAYRN	1,4472	(0,7860. 2,6647)
FCS	FCAYRN	3,43	(1,6621. 7,0785)
FCE	FCBEI	0,3177	(0,1862. 0,5421)
FCHYE	FCBEI	0,3596	(0,1946. 0,6643)
FCS	FCBEI	0,8523	(0,4124. 1,7614)
FCHYE	FCE	1,1318	(0,6746. 1,8988)
FCS	FCE	2,6826	(1,4114. 5,0985)
FCS	FCHYE	2,3701	(1,1577. 4,8521)
Sexo			
Mujer	Hombre	0,6369	(0,4328. 0,9373)
Tipo de plantel			
Publico	Privado	0,442	(0,2745. 0,7115)
Repitió años			
Si	No	2,0688	(1,3141. 3,2569)
Estudios antes			
Si	No	0,6556	(0,4302. 0,9993)
Situación padres			
Separados	Conviven	1,6252	(1,1220. 2,3540)

Fuente: Elaboración propia

Bondad de ajuste del modelo

Para evaluar la significación global del modelo se utiliza el Estadístico Razón de Verosimilitud (ERV), que se define de la siguiente manera:

$$ERV = -2 \left[\ln \frac{L(R)}{L(MV)} \right] = \{-2 \ln[L(R)]\} - \{-2 \ln[L(MV)]\} \quad (12)$$

Donde L(R) es la función de verosimilitud para el modelo restringido en el que únicamente se considera la constante; y L(MV) es la función de verosimilitud para el modelo formulado.

Este estadístico sigue una distribución Chi-cuadrado con (k-1) grados de libertad, donde K es el número de parámetros incluidos en el modelo formulado y que han sido estimados por máxima

verosimilitud, en donde la hipótesis nula es que todos los coeficientes del modelo estimado, excepto la constante son nulos. Esta hipótesis se rechaza si ERV excede el valor crítico. En la Tabla 9 se muestra el valor de ERV y el nivel (P), ante estos resultados, se puede comprobar que efectivamente, la deserción temprana se relaciona con el conjunto de variables estimado.

Tabla 9. Estadísticos de bondad de ajuste del modelo

-2Ln [L(R)]	-2Ln[L(MV)]	ERV	Test de Hosmer-Lemeshow
786,166	704,140	$\chi^2 = 82,026$ P(0,000)	$\chi^2 = 3,31$ (gl=8) P(0,913)

Fuente: Elaboración propia

Otra prueba de bondad de ajuste es la prueba de Hosmer-Lemeshow, que consiste en comparar las frecuencias observadas y esperadas de eventos y no eventos para evaluar hasta qué punto el modelo se ajusta a los datos (Tabla 10). La prueba de Hosmer-Lemeshow sigue una distribución chi-cuadrado con (n-k) grados de libertad y ($P < 0,05$), que tiene como hipótesis nula: El modelo seleccionado ajusta bien los datos (Hosmer y Lemeshow, 1980). Si el valor p para la prueba de bondad de ajuste es menor que el nivel de significancia elegido, las probabilidades pronosticadas se desvían de las probabilidades observadas de una manera que la distribución binomial no predice.

En la Tabla 9 se muestra el valor del test Hosmer-Lemeshow (3,31) con un nivel ($P = 0,913$), ante estos resultados, se acepta el modelo como bueno.

Tabla 10. Frecuencias observadas y esperadas para la prueba de Hosmer-Lemeshow

Grupo	Estado			
	Retiro definitivo		Vigente	
	Observado	Esperado	Observado	Esperado
1	8	9,8	49	47,2
2	14	13,3	43	43,7
3	16	16,4	42	41,6
4	20	19,1	37	37,9
5	22	22,6	36	35,4
6	27	25,3	30	31,7
7	26	29,1	31	27,9
8	38	33,3	20	24,7
9	36	36,6	21	20,4
10	43	44,6	15	13,4

Fuente: Elaboración propia

Otra medida de bondad de ajuste que, así como el test de Hosmer-Lemeshow intenta juzgar la precisión con la que el modelo estimado se aproxima a los datos observados, es, la tasa de predicciones correctas. En la Tabla 11 se aprecia la clasificación que hace el modelo seleccionado y se verifican los criterios de especificidad y sensibilidad del modelo. El punto de corte para la clasificación fue 0,5. Este punto de corte implica que si la probabilidad de desertar estimada por el modelo es superior a 0,5, se pronostica que el estudiante tiene riesgo de desertar durante los primeros cuatro semestres, o que no lo hace en caso contrario, esto es, que para valores menores a 0,5 se pronostica que el estudiante no deserta durante los primeros cuatro semestres.

Tabla 11. Pronósticos del Modelo

Estado	Vigente	Retiro	% de aciertos
Vigente	247	77	76%
Retiro	115	135	54%
Porcentaje global			67%

Fuente: Elaboración propia

En general, se encontró, que el modelo propuesto tiene una tasa de predicciones correctas del 67%, su nivel de especificidad es del 76%, es decir, clasificó correctamente el 76% de los

estudiantes que permanecen vigentes, y su nivel de sensibilidad es del 54%, es decir, clasificó correctamente el 54% de los estudiantes que presentaron deserción temprana.

CONCLUSIONES Y RECOMENDACIONES

El modelo estimado evidencia que un buen puntaje en la prueba Saber 11, el ser mujer, no haber reprobado años durante el bachillerato, el haber cursado estudios antes y si los padres conviven, disminuye la probabilidad de deserción temprana; asimismo, el haber egresado de un colegio privado aumenta la probabilidad de deserción. De todas las facultades en las que se puede matricular el estudiante, las que mayor riesgo tienen de deserción temprana son la de Ciencias Básicas e Ingeniería y la Facultad de Ciencias de la Salud.

Asimismo, de acuerdo a las pruebas de bondad de ajuste se considera que el modelo se adapta a los datos observados. Sin embargo, el resultado global de predicción es relativamente bajo, lo que indica que se pueden llegar a incluir otras variables que aporten explicación a la probabilidad de deserción temprana.

Esta investigación corrobora que la tasa de deserción temprana en la Universidad de los Llanos es alta y sobre todo, logra identificar qué programas presentan los índices más elevados, con estos resultados, se puede esperar que se refuercen y focalicen los planes anti deserción para que se logre disminuir estas cifras.

Las recomendaciones de política educativa que puede aplicar la Universidad de los Llanos se pueden enfocar en realizar pruebas vocacionales para que el estudiante que ingresa acierte en la elección de programa en el que se desea matricular, del mismo modo, es de vital importancia que al detectar a un estudiante en riesgo de deserción se le brinden todas las ayudas disponibles para que este, bien sea, no deserte o tome una mejor decisión. Asimismo, se recomienda una revisión más detallada de la información que suministran los estudiantes cuando realizan la inscripción para que esta pueda ser utilizada con mayor precisión en la detección de posibles desertores.

Se recomienda también, realizar estudios de deserción temprana para cada programa para identificar las causas que afectan a los distintos grupos de estudiantes todo con el fin de obtener mejores resultados cuando se apliquen, de manera focalizada los distintos planes enfocados a prevenir la deserción. Por ejemplo, se puede considerar realizar refuerzos en áreas como la matemática y física y cursos nivelatorios o de profundización en otras áreas como la química y estadística.

REFERENCIAS

- Acevedo D, Torres JD, Tirado DF. 2015. Análisis de la Deserción Estudiantil en el Programa Ingeniería de Alimentos de la Universidad de Cartagena durante el Periodo Académico 2009 - 2013. *Formación universitaria*, 8(1): 35-42.
- Acevedo G, Bayona P. 2011. Identificación de factores motivacionales y sociodemográficos de estudiantes desertores de la facultad de psicología de la universidad Pontificia Bolivariana Bucaramanga. *Psicogente*, 14(25): 132-150.
- Aedo S, Pavlov S, Clavero S. 2010. Riesgo relativo y Odds ratio ¿Qué son y cómo se interpretan? *Revista de obstetricia y ginecología*, 5(1): 51-54.
- Alderete AM. 2006. Fundamentos del análisis de regresión logística en la investigación psicológica. *Evaluar*, 6(1): 52-67.
- Amemiya T. 1975. Qualitative Response Models. *Annals of Economic and Social Measurement*, 4(3): 363-372.
- Bean JP. 1981. Student attrition, intentions, and confidence: interaction effects in a path model. Part 1, the 23 variable model. *Annual meeting of the american educational research association*.
- Becker G. 1964. Human Capital: A theoretical and empirical analysis, with special reference to education. Human Capital, 1st ed. (New York: Columbia University Press for the National Bureau of Economic Research)
- Braxton, J.M., Sullivan, A.V., Johnson, R.M. 1997. Appraising Tinto's theory of college student departure. *Higher education: Handbook of theory and research*, 12(0): 107-164.
- Canales A, De los Ríos D. 2018. Factores explicativos de la deserción universitaria. *Calidad En La Educación*, (26): 173-201.
- Cerda J, Vera C, Rada G. 2013. Odds ratio: aspectos teóricos y prácticos. *Revista médica de Chile*, 41(10): 1329-1335.
- Fernández H, Pérez F. 2005. El modelo logístico: una herramienta estadística para evaluar el riesgo de crédito. *Revista Ingenierías Universidad de Medellín*, 4(6): 55-75

- Giovagnoli P. 2002. Determinantes de la deserción y graduación universitaria: una aplicación utilizando modelos de duración”. Documento de Trabajo 37, Universidad Nacional de la Plata.
- Hanushek, E. A. 1986. The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature*, 24, 1141-1177.
- Hausman JA, McFadden D. 1981. Specification Tests for the Multinomial Logit Model. *Econometrica*, 52(5): 1219-1240.
- Himmel KE, 2002. Modelos de análisis de la deserción estudiantil en la educación superior. *Revista Calidad en la educación*, 0(17): 91-108.
- Hosmer DW, Lemeshow S. 2000. Applied Logistic Regression. 2nd ed. John Wiley & Sons, New York, p. 397.
- Hosmer DW, Lemeshow S. 1980. Goodness of fit tests for the multiple logistic regression model. *Communication in Statistics- Theory and Methods*, 9(10): 1043-1069.
- Jadue JG. 2002. Psychological factors that lead to low achievement, failure and dropping out of school. *Estudios pedagógicos*, (28): 193-204.
- Jaramillo A. 2010. El financiamiento de la educación superior en Colombia: retos y tensiones. *Pensamiento Universitario*, 20: 103-127. Asociación Colombiana de Universidades.
- Jennings DE. 1986. Judging inference adequacy in logistic regression. *Journal of The American Statistical Association*, 81: 471-476
- Larroucau T. 2015. Estudio de los factores determinantes de la deserción en el sistema universitario chileno. *Estudios de políticas públicas*, 0(1): 1-23.
- Lassibille G, Navarro L. 1981. Tratamiento econométrico de las variables cualitativas. *Cuadernos de ciencias económicas y empresariales*, 0(8): 49-92.
- Malagón L, Calderón C, Soto, E. 2006. Estudio de la deserción estudiantil de los programas de pregrado de la universidad de los llanos (1998-2004). *Universidad de los Llanos*, enero de 2006.

- Malagón L, Soto L, Eslava P. 2007. La deserción en la Universidad de los Llanos (1998-2005). *Revista Orinoquia*, 11(1): 23-40.
- Melo-Becerra LA, Ramos-Forero JE, Hernández-Santamaría PO. La educación superior en Colombia: situación actual y análisis de eficiencia. *Revista Desarrollo y sociedad*, 78 (March 2017): 59-111.
- Núñez J. 1990. Una clase de modelos lineales binarios de regresión cualitativa. *Estadística Española*, 32(124): 389-400.
- Olave G, Cisneros M, Rojas I. 2013. Deserción universitaria y alfabetización académica. *Educación y educadores*, 16(3): 455-471.
- Patiño-de Peña L, Cardona AM. 2013. Review of some studies on university student dropout in Colombia and Latin America. *Acta Universitaria*, 23(4): 37-46.
- Porto A, Di Gresia L. 2004. Rendimiento de estudiantes universitarios y sus determinantes. *Revista de Economía y Estadística*, 0(1): 93-113.
- Salas M. 1996. La regresión logística. Una aplicación a la demanda de estudios universitarios. *Estadística española*, 38 (141): 193-217.
- Sánchez F, Quirós M, Reverón C, Rodríguez R. 2002. Equidad social en el acceso y permanencia en la universidad pública determinantes y factores asociados. *CEDE Universidad de los Andes*. Bogotá, Colombia. Recuperado de: https://economia.uniandes.edu.co/components/com_booklibrary/ebooks/D2002-16.PDF
- Silva A, Molina M. 2017. Likelihood ratio (razón de verosimilitud): definición y aplicación en Radiología. *Revista Argentina de Radiología*, 81(3): 204-208.
- Small KA, Hsiao C. 1985. Multinomial Logit Specification Tests. *International Economic Review*, 26(3): 619-627.
- Stare J, Maucourt-Boulch D. 2016. Odds Ratio, Hazard Ratio and Relative Risk. *Metodoloski zvezki*, 13(1): 59-67.
- Tinto V. 1975. Dropout in higher education: a theoretical synthesis of recent research. *Review of Educational Research*, 45(1): 89-125.